RACE TO THE TOP ASSESSMENT

Smarter Balanced Assessment Consortium

Year Three Report

U.S. Department of Education
Washington, DC 20202

July 2014

U.S. Department of Education
Arne Duncan
Secretary of Education

Office of the Deputy Secretary
Jim Shelton
Deputy Secretary

Implementation and Support Unit
Patrick Rooney
Acting Director, Policy and Program Implementation

July 2014

This publication is available on the Department's website at http://www2.ed.gov/programs/racetothetop-assessment/performance.html.

On request, this publication is available in alternate formats, such as Braille, large print or compact disc. For more information, please contact the Department's Alternate Format Center at 202-260-0852 or at 202-260-0818.

Notice to Limited English Proficient Persons

If you have difficulty understanding English, you may request language assistance services for Department information that is available to the public. These language assistance services are available free of charge. If you need more information about interpretation or translation services, please call Edward Ohnemus at (202) 260-1447, or email Edward.Ohnemus@ed.gov.

You may also write to:
Edward Ohnemus
U.S. Department of Education
LBJ Education Building
400 Maryland Ave., SW, Room 5C114
Washington, DC 20202

Introduction

About the Race to the Top Assessment program

The Race to the Top Assessment program was authorized as part of the American Recovery and Reinvestment Act of 2009 (ARRA). In September 2010, the U.S. Department of Education (Department) awarded competitive, four-year grants to two consortia of states, the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (Smarter Balanced).[1]

The two consortia are developing comprehensive assessment systems that are valid, support and inform instruction, provide accurate information about what students know and can do, and measure student achievement against standards, including those that are typically hard to measure, designed to ensure that all students gain the knowledge and skills needed to succeed in college and the workplace. The assessment systems must include one or more summative assessment components in mathematics and in English language arts that are administered at least once during the academic year in grades 3 through 8 and at least once in high school; both consortia are also creating a series of diagnostic, formative, or interim tests that will be available for their member states to provide on-going feedback during the school year to inform teaching and learning. The assessment systems must include all students, including English learners and students with disabilities. PARCC and Smarter Balanced will each develop a common measure for use by their member states whether individual students are college and career ready or on track to being college and career ready. The assessment systems will provide an accurate measure of student achievement, particularly for very high- and low-achieving students, and an accurate measure of student growth over a full academic year or course.

These assessment systems, which will be operational in the 2014-2015 school year, are intended to play a critical support role in educational systems; provide administrators, educators, parents, and students with the data and information needed to continuously improve teaching and learning; and help meet the President's goal of restoring, by 2020, the nation's position as the world leader in college graduates.

Race to the Top Assessment Program Review

As part of the Department's commitment to supporting states as they implement ambitious reform, the Department established the Implementation and Support Unit (ISU) in the Office of the Deputy Secretary to administer, among others, the Race to the Top Assessment program. The goal of the ISU is to provide collaborative support to grantees as they implement unprecedented and comprehensive reforms to improve student outcomes. Consistent with this goal, the Department has developed a Race to the Top Assessment program review process that not only addresses the Department's responsibilities for fiscal and programmatic oversight, but is designed to identify areas in which the consortia need assistance and support to meet their goals. The ISU works with the Race to the Top Assessment consortia to identify and provide support based on their specific plans and needs. ISU staff encourages collaboration and partnership across the consortia and with outside experts to achieve and sustain educational reforms that improve student outcomes.

The consortia are accountable for implementing their approved Race to the Top Assessment plans. The program review is a continuous improvement process.[2] Regular updates and data from the consortium inform the Department's support for the consortia. In the event that adjustments are required to an approved plan, the consortium must submit a formal amendment request to the Department for consideration. The consortia may submit for Department approval amendment requests to a plan and

[1] More information about the Race to the Top Assessment program is available at www.ed.gov/programs/racetothetop-assessment.

[2] More information about the ISU's Race to the Top Assessment program review process can be found at: www.ed.gov/programs/racetothetop-assessment/review-guide.pdf.

budget provided that such changes do not significantly affect the scope or objectives of the approved plans. The consortia's approved plans, including any approved amendments, can be found at: www2.ed.gov/programs/racetothetop-assessment/awards.html.

If the Department determines that the consortium is not meeting its goals, activities, timelines, budget, or annual targets or is not fulfilling other applicable requirements, the Department will take appropriate enforcement action(s), consistent with 34 CFR § 80.43 in the Education Department General Administrative Regulations (EDGAR).

ABOUT THIS REPORT

This report on the consortium's Year 3 implementation of the Race to the Top Assessment program is focused on the four primary components of the consortium's activities: governance; assessment design and development; professional capacity, outreach, and communications; and technology. The Department used the information gathered during the program review process (e.g., through monthly calls, an on-site visit conducted in December 2013, the consortium's annual performance report (APR) which was submitted in September 2013, and the Department's Technical Review in spring 2013) to draft the report.

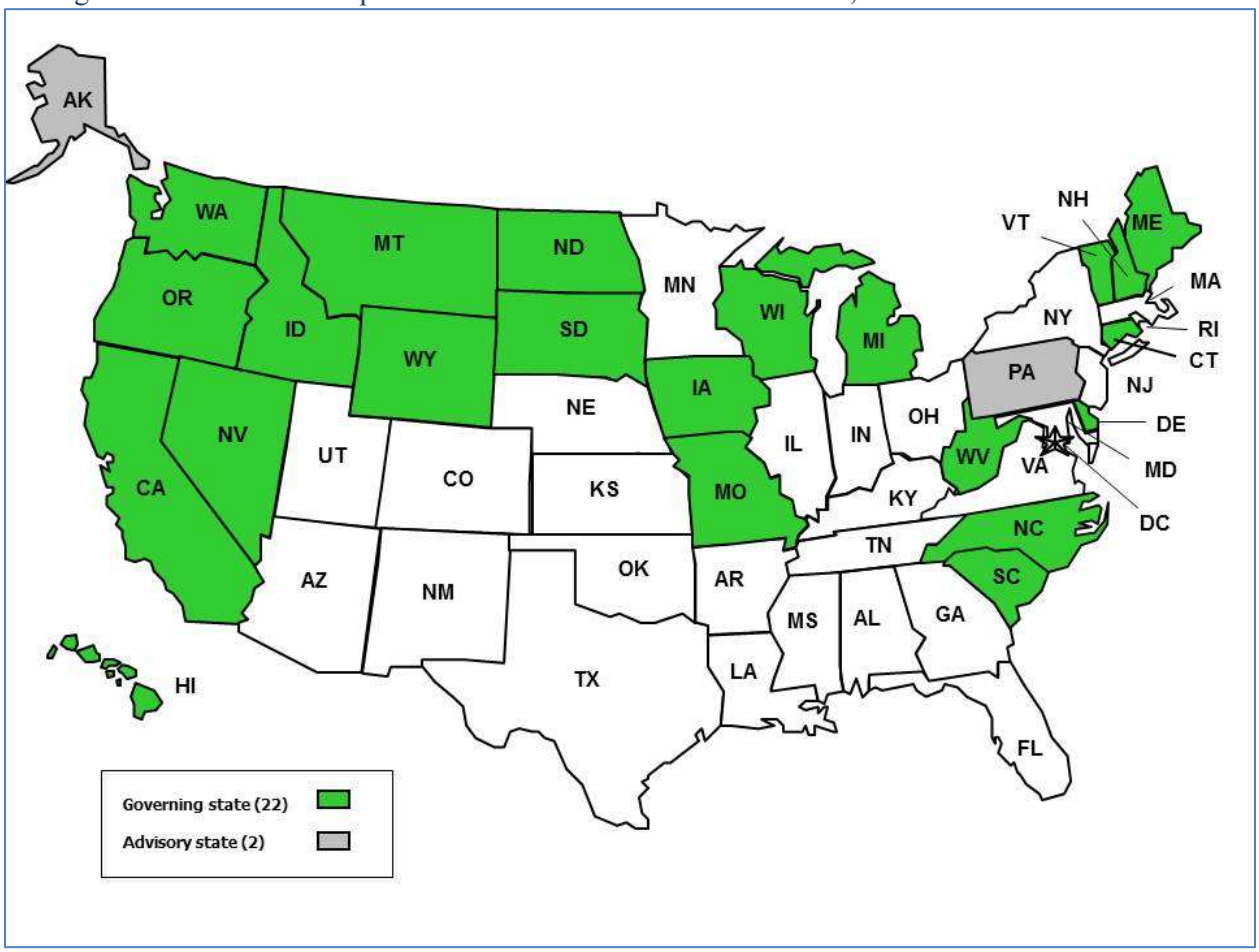
This report serves as an assessment of the consortium's overall implementation of its approved plan, highlighting successes and accomplishments, identifying challenges, and noting important lessons learned during the third year and key upcoming activities. It covers the period from January 2013 through the end of December 2013, except in a few instances where more recent information is noted.

About the Smarter Balanced Assessment Consortium

The Smarter Balanced Assessment Consortium (Smarter Balanced) is designing a computer-adaptive assessment system to measure whether students are ready for college or career or on track to being college and career ready. The assessment system will include summative assessments, optional interim assessments, and a digital library consisting of formative tools and resources to support educators.

As of December 31, 2013, Smarter Balanced consists of 24 states (see figure 1). Twenty-two are governing states, meaning they are involved in policy decision-making for the consortium, and two are advisory states, meaning they are involved in the work of Smarter Balanced but are less actively engaged in determining consortium policies. During Year 3, North Dakota and Wyoming shifted from being advisory to governing states and one state, Kansas, left the consortium. In addition, the U.S. Virgin Islands joined the consortium as an affiliate member, partaking in the consortium's activities at its own expense. (Only the 50 States, the District of Columbia, and Puerto Rico are eligible to be governing or advisory states per the requirements of the American Recovery and Reinvestment Act). Awarded a grant in the amount of $175,649,539 by the Department in September 2010, Smarter Balanced selected Washington to serve as its fiscal agent. The consortium has contracted with WestEd as its project management partner.

Figure 1. State membership in Smarter Balanced as of December 31, 2013



NOTE: Alaska withdrew on January 14, 2014, and South Carolina withdrew on April 14, 2014, bringing the total, as of May 31, 2014, to 21 governing states, 1 advisory state, and 1 affiliate member.

THEORY OF ACTION

The Smarter Balanced application included a theory of action based on several principles of assessment systems in high-achieving nations and states:[3]

1. Assessments are grounded in a thoughtful, standards-based curriculum and are managed as part of an integrated system of standards, curriculum, assessment, instruction, and teacher development. Curriculum and assessments are organized around a well-defined set of learning progressions along multiple dimensions within subject areas. Formative and interim/benchmark assessments and instructional supports are conceptualized in tandem with summative assessments—all of them linked to the standards and supported by a unified technology platform.
2. Assessments produce evidence of student performance on challenging tasks that evaluate the Common Core State Standards (CCSS). Instruction and assessments seek to teach and evaluate knowledge and skills that generalize and can transfer to higher education and multiple work domains. They emphasize deep knowledge of core concepts and ideas within and across the disciplines—along with analysis, synthesis, problem solving, communication, and critical thinking—thereby requiring a focus on complex performances as well as on specific concepts, facts, and skills.
3. Teachers are integrally involved in the development and scoring of assessments. While many assessment components are efficiently scored with computer assistance, teachers must also be involved in the formative and summative assessment systems so that they deeply understand and can teach in a manner that is consistent with the full intent of the standards, while becoming more skilled in their own assessment practices.
4. The development and implementation of the assessment system is a state-led effort with a transparent and inclusive governance structure. Starting in December 2009, prior to being awarded an RTTA grant, Smarter Balanced has hosted weekly conference calls and several face-to-face meetings open to all states interested in establishing a consortium of states for the development of assessments aligned to the CCSS. Those activities have resulted in a governance structure that has established a consensus decision-making model and clear leadership roles. Each state's commitment to the collaborative process and products will facilitate the development of a complex system and signal ongoing support for its implementation.
5. Assessments are structured to continuously improve teaching and learning. Assessment as, of, and for learning is designed to develop understanding of what learning standards are, what high-quality work looks like, what growth is occurring, and what is needed for student learning.
6. Assessment, reporting, and accountability systems provide useful information on multiple measures that is educative for all stakeholders. Reporting of assessment results is timely and meaningful—offering specific information about areas of performance so that teachers can follow up with targeted instruction, students can better target their own efforts, and administrators and policymakers can more fully understand what students know and can do, in order to guide curriculum and professional development decisions.
7. Design and implementation strategies adhere to established professional standards. The development of an integrated, balanced assessment system is an enormous undertaking, requiring commitment to established quality standards in order for the system to be credible, fair, and technically sound.

ASSESSMENT SYSTEM DESIGN

As Smarter Balanced described in its application, it will develop an assessment system that "promotes research-supported instructional practice and incorporates a balanced set of technology-supported tools, innovative assessments, and state-of-the-art classroom support mechanisms that work coherently to

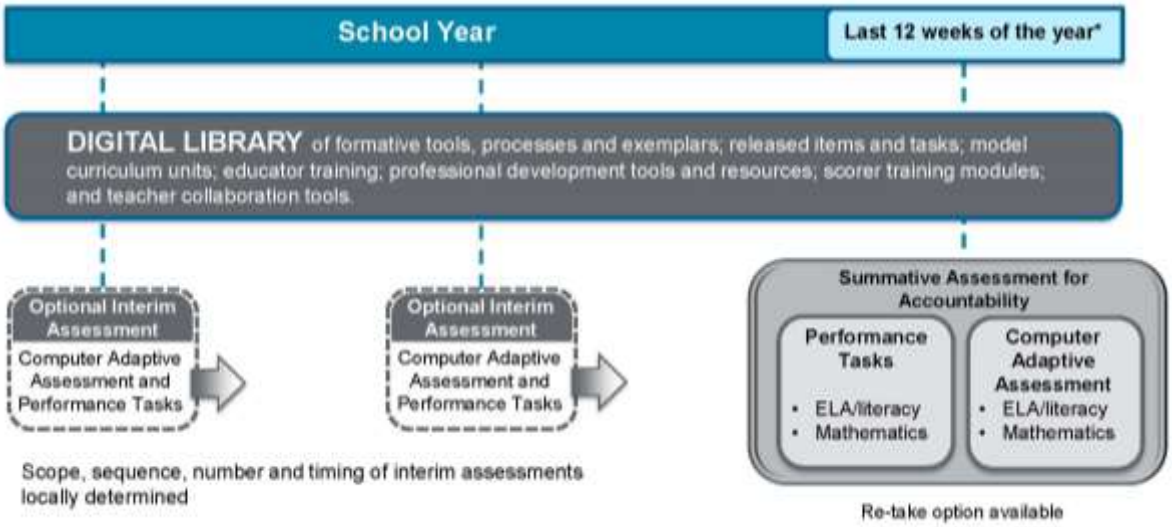
[3] See pages 32-33 of the consortium's application, available at www2.ed.gov/programs/racetothetop-assessment/rtta2010smarterbalanced.pdf.

support teaching and learning" (p. 6 of the application). The assessment system will be comprised of formative, interim, and summative components.

Figure 2 details the primary components of the assessment system. Specifically, as noted in the application, Smarter Balanced proposes to implement a system that features the following (p. 5-6):

- Computer-adaptive summative assessments that make use of technology-enhanced item types and teacher-developed and scored performance events;
- Computer-adaptive interim/benchmark assessments – reflecting learning progressions or content clusters – that provide more in-depth and/or mid-course information about what students know and can do;
- Research-supported instructionally sensitive tools, processes, and practices developed by state educators that can be used formatively at the classroom level to improve teaching and learning;
- Focused on-going support to teachers through professional development opportunities and exemplary instructional materials;
- Online reporting and tracking system that enables access to key types of information about student progress toward college- and career-readiness and about specific strengths and limitations in what students know and are able to do at each grade level; and
- Cross-state communications network to inform stakeholders about Smarter Balanced activities and ensure a common focus on the goal of college- and career-readiness for all students.

Figure 2. Smarter Balanced assessment system



* Time windows may be adjusted based on results from the research agenda and final implementation decisions.
Source: www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Smarter-Balanced-Overview-Presentation.pdf.

Smarter Balanced will develop summative assessments for grades 3 through 8 and 11 with optional assessments available for grades 9 and 10. The Smarter Balanced summative assessments will consist of:

- Performance tasks that require student-initiated planning, management of information and ideas, interaction with other materials or people, and production of more extended responses (p. 53 of the application).
- Computer-adaptive assessment that selects items for a student based on the student's responses to previous questions, thus adapting to the student's demonstrated ability throughout the test. This component will include selected response, constructed response, and technology-enhanced

constructed response items. It will be administered during the last 12 weeks of the school year. Students will have an opportunity to re-take this component if necessary.

The assessment system will be computer-administered. Smarter Balanced anticipates a three-year phase-in period (until the 2017-2018 school year) during which school districts may administer a comparable paper-and-pencil version of the summative assessment.

In Year 3, the consortium amended its plan for translating the mathematics assessment so that it will provide a stacked Spanish translation for English learners enrolled in Spanish/English dual-language programs and will provide pop-up glossaries in mathematics for the nine most commonly spoken languages and dialects in the Consortium other than English among English learners in member states (Spanish, Vietnamese, Arabic, Filipino, Cantonese, Mandarin, Punjabi, Korean, Ukrainian and Russian). English language arts listening items, as well as mathematics items and performance tasks, will be translated into American Sign Language (ASL).

The Smarter Balanced assessment system will also include interim assessments that will be available to member states and a digital library that includes formative tools, processes, and practices for teachers to use to support instruction. During Year 3, Smarter Balanced refined its approach to providing optional interim assessments in grades 3-11. Smarter Balanced will provide two types of interim assessments for member states. The first is an "interim comprehensive assessment," which uses the same blueprint as the summative assessment. The second is comprised of "interim assessment blocks," which focus on blocks of items that provide targeted information about discrete knowledge and skills to aid instruction.

More information about Smarter Balanced can be found at: www.smarterbalanced.org.

Assessment Design and Development

The extent to which the consortium is developing a comprehensive assessment system that measures student knowledge against the full range of the college- and career-ready standards, including the standards against which student achievement has traditionally been difficult to measure; provides an accurate measure of achievement, including for high- and low-performing students, and an accurate measure of student growth over a full academic year or course; and produces student achievement data and student growth data that can be used to determine whether individual students are college- and career-ready or on track to being college- and career-ready.

PILOT TEST & ITEM DEVELOPMENT

In Years 1 and 2 of the grant, Smarter Balanced laid the groundwork for the initial assessment development by creating English language arts and mathematics content specifications that provided the initial assessment targets on which the assessments were based. Among other things, the specifications identify the "claims" about student learning that will be measured and reported on the summative assessment:

English language arts

- Overall English language arts and literacy performance.
- Read closely and analytically to comprehend a range of increasingly complex literary and informational texts.
- Produce effective and well-grounded writing for a range of purposes and audiences.
- Employ effective speaking and listening skills for a range of purposes and audiences.
- Engage in research/inquiry to investigate topics and to analyze, integrate, and present information.

Mathematics

- Overall mathematics performance.
- Concepts and procedures – Explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.
- Problem solving & modeling and data analysis – Solve a range of complex well-posed problems in pure and applied mathematics, making productive use of knowledge and problem-solving strategies. Analyze complex, real-world scenarios and construct and use mathematical models to interpret and solve problems.
- Communicating reasoning – Clearly and precisely construct viable arguments to support the student's own reasoning and to critique the reasoning of others.

Smarter Balanced also developed a wealth of documents (which are publicly available on the consortium's website) to guide the writing and reviewing of items and tasks. The consortium used these resources to develop approximately 5,300 English language arts and 4,900 mathematics items and tasks across grades 3-8 and 11. The items included selected-response, constructed-response, and technology-enhanced items, as well as performance tasks. Items were reviewed by the Smarter Balanced English language arts and mathematics content directors, the consortium director of support for underrepresented students, state staff who are consortium work group members, and educator committees made up of experts from member states in academic content, bias and sensitivity, and accessibility for students with disabilities and English learners. Over 160 educators were involved in writing and reviewing the items and tasks developed for the pilot test. Higher education representatives and external content experts also reviewed grade 11 items.

In spring 2013, Smarter Balanced conducted a computer-administered pilot test to over 650,000 students in more than 5,000 schools in the consortium's governing states. Smarter Balanced gathered essential information from the pilot test about the accessibility features of the assessment system, the performance of individual test items, and the functionality of the test platform. As part of the pilot test, Smarter Balanced conducted a survey of participating students and educators.

The work to prepare for and conduct the pilot test informed the consortium's work developing additional items in Year 3 for the field test that will be conducted in spring 2014. Several challenges were noted by Smarter Balanced during item development, particularly regarding the quality control measures. The Department's Technical Review in spring 2013, which brought together national assessment experts to review the consortium's progress, also highlighted the need for Smarter Balanced to review and revise its item development process prior to the next phase of item development.[4]

Following the pilot test, in Year 3, Smarter Balanced developed approximately 21,000 items and tasks in English language arts and mathematics in advance of the field test that will be conducted in spring 2014. Smarter Balanced made several changes to the item development process to address concerns about the quality of the items developed for the pilot test. The consortium established an item quality review panel (IQRP) consisting of experts in English language arts, mathematics, creating accessible items, and assessment design. The IQRP reviewed and revised the item development documents and established a set of quality criteria to guide the development of items and tasks. The panel assisted in the development of 250 item and task archetypes in each content area to serve as exemplar items for item authors. To help manage the scale of the project and the short timelines while maintaining quality, Smarter Balanced also improved the criteria for the review and approval of items. For example, item authors were required to submit a set of sample items for review and approval by the consortium and IQRP before receiving approval to write additional items. As was done for the pilot test, field test items and tasks were written by trained item authors, including approximately 300 educators. All field test items went through multiple layers of review by tiers of content experts that were comprised of contractors Smarter Balanced staff and state educators, and higher education representatives, to review the content, ensure it was free of bias and sensitivity, and that it was accessible to all students.

The consortium also included a pilot of state-managed item writing, where states volunteered to lead efforts to write items for the consortium. Eight states each contributed approximately 300 items. Following the field test, Smarter Balanced will evaluate all items, including the state-managed items, to determine their quality and appropriateness for either the summative or interim item pools. The consortium anticipates that state-managed items will be used in the future to help expand the interim assessment pool.

Sample Items and Practice Tests

In Year 2, Smarter Balanced released sample items online that provided states, stakeholders, and the public with an early view of the types of items likely to be on the Smarter Balanced assessment system. In Year 3, following the conclusion of the pilot test in May 2013, the consortium released practice tests for English language arts and mathematics for grades 3-8 and 11. Access to the practice test is available to all and gives teachers, administrators, parents, and students a preview of a variety of Smarter Balanced item types, as well as an opportunity to try out the various accommodations and supports that are available for students, such as text-to-speech, Braille, ASL, and the mathematics pop-up glossaries.

Achievement Level Descriptors

Achievement level descriptors (ALDs) are statements that articulate the knowledge, skills, and abilities represented at different levels of performance on the Smarter Balanced assessments. They describe how

[4] More information about the Technical Review can be found at: www2.ed.gov/programs/racetothetop-assessment/performance.html.

students are progressing toward mastery of the CCSS and provide clear explanations of student performance for educators, parents, and policymakers. In October 2012, based on collaboration among the Smarter Balanced workgroups, K-12 state leads, higher education leads, and state educators, Smarter Balanced released initial ALDs for grades 3-8 and 11 for public comment. Following the public comment, the initial ALDs were adopted by the governing states in March 2013.[5] After achievement standards are set in fall 2014, the consortium will review the ALDs and revise them, if necessary.

As part of the development of the ALDs, the consortium also focused on how it will measure a student's college readiness. Due to the wide array of factors involved in whether a student is college ready, Smarter Balanced focused on whether students are college content ready. In April 2013, the Smarter Balanced member states adopted an initial college content readiness policy that defines readiness for each content area:

> *English language arts*
> Students who perform at the College Content-Ready level in English language arts/literacy demonstrate reading, writing, listening, and research skills necessary for introductory courses in a variety of disciplines. They also demonstrate subject-area knowledge and skill associated with readiness for entry-level, transferable, credit-bearing English and composition courses.
>
> *Mathematics*
> Students who perform at the College Content-Ready level in mathematics demonstrate foundational mathematical knowledge and quantitative reasoning skills necessary for introductory courses in a variety of disciplines. They also demonstrate subject-area knowledge and skills associated with readiness for entry-level, transferable, credit-bearing mathematics and statistics courses.

ACCESSIBILITY

During Year 3, Smarter Balanced took several important steps toward making sure the assessment system is accessible for students with disabilities and English learners. In September 2013, the governing states approved the *Usability, Accessibility, and Accommodations Guidelines*. This research-based policy, which was approved after a public comment period, outlines three categories of resources that will be available to students:

1. Universal tools, such as a digital notepad, will be available to all students.
2. Designated supports, such as a translated pop-up glossary, will be available to any student, provided someone familiar the student's needs and testing resources indicates that it is necessary for the student.
3. Accommodations, such as Braille and ASL, will be available to students with a documented need noted in an Individualized Education Program (IEP) or 504 plan.

For the field test, the consortium will ask participating schools to identify accommodations or designated supports needed by the student that are not permitted by the *Guidelines*. Following the conclusion of the field test, Smarter Balanced will evaluate each request to determine whether it should be permitted for the operational assessment in the 2014-2015 school year and will revise the *Guidelines* as necessary.

Based on the pilot test experience in spring 2013 and consultation with experts in how to assess English learners, the consortium requested, and the Department approved, a change in the approach to translating assessments. The Smarter Balanced application contemplated translating the mathematics assessment into five languages. Based on the data from the pilot test and research from the field, providing a pop-up glossary was more commonly used by English learners and as helpful as a full translation. Smarter Balanced will provide a full Spanish stacked translation for those students enrolled in a dual-language

[5] To view the complete ALDs and college content readiness policy, see www.smarterbalanced.org/achievement-level-descriptors-and-college-readiness/.

immersion program. For all other English learners, Smarter Balanced will provide pop-up glossaries in mathematics for ten languages: Spanish, Vietnamese, Arabic, Filipino, Cantonese, Mandarin, Punjabi, Korean, Ukrainian, and Russian. In addition, the glossaries will be able to account for various dialects of the languages (e.g., between Spanish-speaking students of Mexican, Salvadoran, or Puerto Rican descent). The consortium will provide ASL for English language arts listening items and the mathematics assessments.

RESEARCH

Smarter Balanced continues to rely on its technical advisory committee (TAC) for guidance and support as it researches the validity, reliability, fairness, and technical quality of its assessment system. During Year 3, Smarter Balanced convened the TAC for two in-person meetings and five webinars. The TAC provided feedback on the consortium's overall validity research plan; the development of ALDs; the career readiness definition; the approach to working with PARCC on studying comparability between the assessment systems; the design for setting achievement standards; the test blueprints; and the design for both the pilot and field tests and related research studies. The TAC also provided feedback to Smarter Balanced on the question of what effect the testing window might have on student performance and how to include external measures on the field test as a benchmarking study, such as National Assessment of Educational Progress (NAEP), Program for International Student Assessment (PISA), PARCC, and the SAT.

During Year 3, the consortium crafted a vision for the use of automated scoring in the assessment system and conducted several analyses, including during the pilot test in spring 2013. Based on the lessons learned, and after discussions with the TAC, Smarter Balanced worked to develop more specific scoring rubrics to improve consistency in scoring. Other analyses conducted on the pilot test data included an analysis of dimensionality, item response theory (IRT), scaling, calibration, and differential item functioning (DIF), an evaluation of the accommodations and translations, and an analysis of whether it is helpful to include a classroom-based component on the performance tasks to ensure a basic level of understanding of the context and construct-irrelevant vocabulary of performance task stimulus materials.

Smarter Balanced continued conducting cognitive labs in Year 3. The labs helped Smarter Balanced understand how students work through and solve various types of items by having students discuss their thinking while answering test items. The consortium reported that it gathered valuable information around the use of item formats to measure certain cognitive skills; how well the use of pop-up glossaries helped improve the validity of the score for English learners.; and how much time it takes to answer items in different formats. During the third year of the grant, Smarter Balanced also continued developing the algorithm that will guide the computer-adaptive portion of the summative assessments. The consortium conducted simulations and consulted with its TAC as it defined the algorithm components, the goals of the item selection algorithm, the core design principles, and the item selection method.

During the pilot test in spring 2013, Smarter Balanced explored the feasibility of scoring hand-scored items through the use of automated scoring software. This involved double-scoring (with expert resolution when the two readers disagree) many pilot test items, representing the different types of items. Software developers then tried to replicate the hand-scored results. Findings from this work on the pilot test will contribute to refinements of this strategy and additional research for the spring 2014 field test.

Finally, Smarter Balanced developed plans for evaluating the alignment of the assessment system to the college- and career-ready content standards in Year 3. The TAC provided feedback and endorsed the Smarter Balanced approach to collecting evidence of alignment. The alignment study will begin in spring 2014; some components of the study cannot be undertaken until after the first operational administration in the 2014-2015 school year.

LESSONS LEARNED

From the pilot test, Smarter Balanced learned several important lessons about the initial items that were developed. For example, survey responses from the participants indicated that many of the items and tasks were more difficult than the consortium anticipated. Data indicated that some item types were not appropriate for the content or skill addressed in the assessment target. In addition, the pilot test helped identify the need to change particular types of items to improve the accessibility and testing experience and to revise the administration manuals and training materials. The Department's Technical Review highlighted the need for Smarter Balanced to review and revise its item development process to include additional quality controls. As discussed above, the consortium substantially revised its development process. In addition to establishing the IQRP, Smarter Balanced reviewed and revised the materials that support item development, such as the content specifications and item writer training materials. The consortium increased the number of reviews and quality checks in the item development process. After conducting research and consulting with the TAC, Smarter Balanced revised the number of items to be developed for the field test (from 38,000 to approximately 21,000). While still developing sufficient items to cover the full range of the standards, the consortium reduced the items to acknowledge the tight timelines and need to focus on the quality of the items that are developed. (Smarter Balanced will develop additional items in Year 4.) Even with this change, the consortium faced challenges developing items in time for the field test that met the consortium's quality criteria. Consortium leadership was actively engaged in identifying and trying to resolve problems throughout fall 2013 so that the consortium met its goal for the number of high-quality items to be developed for the field test.

Additionally, as a result of the lessons learned from the pilot test development and the Department's Technical Review, the consortium developed a text complexity guide to assist item authors in developing passage and item sets. The consortium also increased the percentage of reading passages that are based on published (versus commissioned) from 45 percent for the pilot test to 70 percent for the field test. Smarter Balanced contracted with the Copyright Clearance Center to streamline the process to receive permission to use specific passages. Finally, the consortium revised the assessment blueprints to include one additional reading passage in grades 3-5 and two additional reading passages in grades 6-8 and 11. The consortium also determined that including a classroom-based component prior to having the students take the performance task will help ensure students have a necessary level of understanding of the contextual information and construct-irrelevant vocabulary included in performance task stimulus materials.

LOOKING AHEAD

The major focus of the consortium during Year 4 of the grant will be the administration and scoring of the field test. This will be a significant test of the Smarter Balanced assessment system. It will permit the consortium to evaluate the items and tasks that have been developed and which will comprise the item pool for the system when it becomes operational in the 2014-2015 school year. The field test will also provide an opportunity to test the assessment system with a significant number of students to evaluate the test delivery system, accommodations, designated supports, and accessibility tools to identify and correct problems before the test becomes operational. The field test will be computer administered and will take place from mid-February to early June. More than three million students across the member states will participate. At least 20 percent of all governing states will participate in the field test and nearly all students in California, Connecticut, Idaho, Montana, and South Dakota will participate, providing a good test of the consortium's infrastructure to deliver the computer-administered field test.

Data collected from the field test will be used to inform future item writing, evaluate accessibility features and accommodations, and determine whether changes are necessary to the test administration instructions and manuals. In addition, Smarter Balanced will conduct research, including examining the quality of the items that have been developed to date and studying the comparability of the computer-administered and paper-and-pencil forms of the assessment. Smarter Balanced will also continue to evaluate the design of the computer-adaptive test (CAT) component of the assessment system to ensure that it is appropriately

measuring the full range of the content standards while tailoring the difficulty of the assessment to the student's knowledge and skills. The field test will also yield valuable information about the reasonableness of using automated scoring strategies on items that are traditionally scored by hand.

Based on the data from the field test, Smarter Balanced will set initial achievement standards in fall 2014. The process will consist of several phases. In phase I, up to 250,000 educators, higher education faculty, parents, and other interested individuals from all member states will be able to review items and data from the field test using an online tool to make recommendations for the Level 3 cut score.[6] In phase II, Smarter Balanced will recruit panels of educators and other stakeholders to participate in an in-person meeting to recommend cut scores. In phase III, the consortium will establish a vertical articulation committee, from individuals who participated in phase II, which will use data from phases I and II to ensure that the recommended cut scores are aligned across grades 3-8 and 11. These recommendations will be brought to the member states for review and approval. Smarter Balanced will review the initial achievement standards following the operational assessment in the 2014-2015 school year.

A third phase of item development will take place in spring and summer 2014 to provide approximately 11,000 additional items to be field tested during the operational test in spring 2015 to help strengthen the item pool and aid the sustainability of the consortium. Smarter Balanced will also conduct additional cognitive labs in Year 4.

In Year 3, Smarter Balanced established the Career Readiness Task Force to describe student performance on the grade 11 summative assessments in relation to academic readiness for postsecondary career education and training. The task force consists of individuals from state education agencies, community colleges, institutions of higher education, and career and technical education schools, and other experts. The task force has been drafting a career readiness policy. In Year 4, the career readiness policy will be distributed for public comment before being voted on by member states in spring 2014.[7]

[6] The Level 3 cut score will be the cut score that demonstrates a student is college and career ready.
[7] Smarter Balanced governing states approved the Career Readiness Frameworks in April 2014. More information is available at: www.smarterbalanced.org/wordpress/wp-content/uploads/2014/02/Understanding-and-Using-the-Career-Readiness-Frameworks-V2-Public-Review.pdf.

Professional Capacity, Outreach, and Communications

The extent to which the consortium is supporting member states in implementing rigorous college- and career-ready standards, supporting educators in implementing the assessment system, and informing and building support among the public and key stakeholders.

Educators have been and continue to be heavily engaged in the consortium's work. More than 450 educators have been involved in writing and reviewing the items and tasks developed by Smarter Balanced for the pilot and field tests and many more are engaged in building the Smarter Balanced digital library.

PROFESSIONAL CAPACITY

In Year 3, Smarter Balanced began developing a digital library consisting of formative assessment practices and professional learning resources for educators. Once launched, it will provide educators access to classroom-based, formative assessment strategies and practices that enhance day-to-day instruction and resources to interpret and make use of the data and reports. Through the digital library, the consortium will make available modules that help educators improve their assessment literacy. The digital library resources will also include exemplar instructional modules and additional resources from external sources. Educators will have the capability to review, download, and rate resources; search to find resources by specific content standards and other topics; collaborate and share their knowledge; and access resources that are stored in participating libraries.

In Year 3, Smarter Balanced formed three groups of educators to drive its work developing resources for the digital library:

- The Formative Assessment Advisory Panel (FAAP) is comprised of 11 experts in formative assessment, classroom instruction, gifted education, English language arts, mathematics, online learning, rural and urban education, and supports for English learners and students with disabilities. The FAAP developed and recommended criteria for evaluating the quality of consortium-developed and other educational resources and for approving them for the library.
- The State Leadership Teams (SLTs) are comprised of 8-12 educators per state, including K-12 educators and higher educational faculty. They recruit and train educators for the State Networks of Educators, monitor and support their work, help review resources, and help make final publishing decisions.
- The State Networks of Educators (SNEs) consist of between 70 and 150 educators per state. SNEs are comprised of K-12 educators and higher education faculty with expertise in English language arts, mathematics, science, social studies, gifted and talented, English learners, and students with disabilities. After being trained on the quality criteria established by the FAAP, the SNEs review resources, identify additional resources for review, serve as early users of the system, and provide feedback on the library resources, the review and publishing process, the quality criteria, and the usability of the library.

By December 2013, both SLTs and SNEs had been trained to evaluate the quality of the resources. A rubric and evaluation criteria developed by the SLTs is used by both groups to ensure that the resources reflect the intent of the CCSS, reflect the Smarter Balanced vision of effective formative assessment practices, and are of the high quality expected by the consortium. Due to delays in getting this work started, the digital library will now be launched in several waves in Year 4. Smarter Balanced will continue to add resources to the library in Year 4 and after the grant ends.

In December 2013, the consortium received a grant from the Helmsley Charitable Trust to work with the American Federation of Teachers (AFT) and National Education Association (NEA). The grant will increase the number of educators serving on the SNEs, adding AFT and NEA members from each

member state, to identify instructional resources and develop high-quality exemplar instructional and professional development modules. In addition, the grant will permit the consortium to collaborate with 125 "teacher ambassadors" to provide assessment training to more than 1,200 teachers across member states.

COMMUNICATIONS

Communicating regularly and effectively across a large number of member states to such a diverse population and set of interested stakeholders continues to present challenges for Smarter Balanced. Clear, consistent communication is required to ensure that state, district, and school personnel, as well as key stakeholders, understand the consortium's progress and activities, support the consortium's assessment development, and know what is expected of them to further the consortium's success and sustainability.

The primary vehicle for communicating information about the consortium remains the public website, www.smarterbalanced.org. The website provides a wealth of information about the consortium's progress, technical documentation related to the development of the assessments, and information about the consortium for parents and other interested groups. In Year 3, Smarter Balanced improved the website to provide key information about the consortium in Spanish. In addition to the public website, Smarter Balanced maintained its internal, password-protected website for member states. This site permits the consortium to share documents for internal review before they are made public and eases communication across the geographically diverse member states. In addition to the public website, in October the consortium launched a weekly notification to its state leads that summarizes upcoming events, projects that are underway, and actions taken by the executive committee or the governing states.

Smarter Balanced continued to provide opportunities for the chief state school officers of its governing states to engage in the decision-making structure of the consortium through regularly scheduled webinars or teleconferences and during two all-states collaboration conferences. In addition, the consortium takes advantage of other opportunities to meet with the chief state school officers are meeting, such as during in-person meetings convened by the Council of Chief State School Officers. In Year 3, Smarter Balanced also held several regional leadership meetings. Each was attended by state teams composed of chief state school officers and the state higher education leaders.

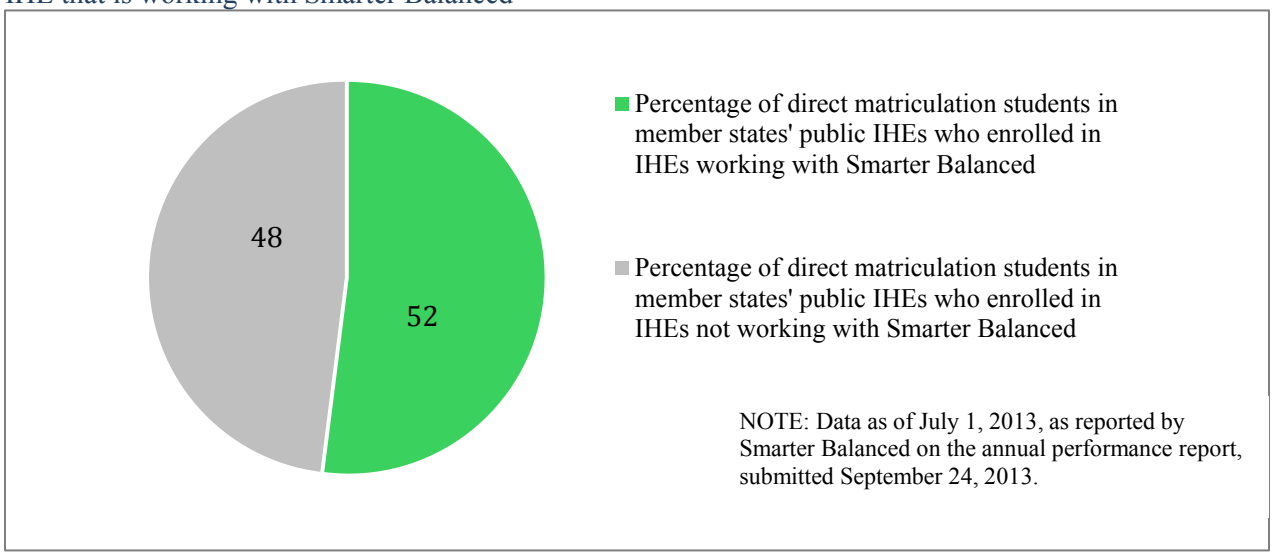
The consortium also continued to produce a quarterly newsletter and weekly email for partners and stakeholders. Smarter Balanced also established several partnerships in Year 3. For example, the National Parent Teacher Association (PTA) developed state-specific information regarding assessments in each state that has adopted the CCSS which are publicly available at www.pta.org/advocacy/content.cfm?ItemNumber=3816.

HIGHER EDUCATION ENGAGEMENT

Smarter Balanced continued to work with higher education in Year 3. Each member state has identified a representative to serve as the higher education state lead, such as higher education leaders and faculty members. Smarter Balanced continues to use five regional advisors to provide guidance and assistance to the higher education state leads. The advisors visit their states in the region at least twice each year and meet with the state leads by phone monthly. They work with the states in their region to assist with the state's implementation plan to build support for the consortium and establish policies for exempting from remedial courses those students who meet the consortium's standards for being college and career ready. Institutions of higher education (IHEs) were actively engaged in defining the ALDs, aligning the ALDs in grades 8 and 11, defining college readiness and developing the definition of career readiness, and developing and reviewing grade 11 items and tasks.

Figure 3 provides data from the Smarter Balanced annual performance report (APR) regarding the involvement of higher education involvement in the consortium. As of July 1, 2013, Smarter Balanced is working with 123 IHEs[8] that have committed to implementing policies that exempt from remedial courses and place into credit-bearing college courses any student who meets the consortium-adopted achievement standard for each assessment and any other placement requirement established by the IHE or IHE system.

Figure 3. Percentage of direct matriculation students in Smarter Balanced member states who attend an IHE that is working with Smarter Balanced



Lessons Learned

The consortium continues to work on methods to expand its outreach efforts. Smarter Balanced continues to make a wealth of materials available on its website and distribute monthly and quarterly updates on the consortium's progress. During Year 3, the consortium also developed a Spanish section of its website.

Over the past year, the consortium identified challenges with the work on the digital library, which led to the delay of the launch of the library until Year 4. Identifying and developing resources that meet the consortium's quality criteria proved challenging. Smarter Balanced developed quality criteria and tools to help SLT and SNE members identify appropriate resources for the digital library. Smarter Balanced is committing to developing a digital library with high-quality tools and resources to support educators

Looking Ahead

In Year 4, Smarter Balanced will launch the digital library in several waves and continue working with educators, through the SLTs and SNEs, to develop and identify resources to be housed in the digital library. In addition, the work initiated due to the grant from the Helmsley Charitable Trust will begin with the AFT and NEA to identify digital library resources and developing teacher ambassadors. Smarter Balanced will continue its efforts to engage higher education in developing the consortium's assessment system, including their involvement in setting achievement standards.

[8] The number of IHEs working with Smarter Balanced has changed during the grant period due to states having left the consortium. States and others that have joined the consortium since the grant was awarded (California, Vermont, and the U.S. Virgin Islands) have not submitted letters of intent for their IHEs.

Technology

The extent to which the consortium is using technology to the maximum extent appropriate to develop, administer, and score assessments and report results.

In Year 3, the consortium made important strides developing the technology system necessary to carry out the Smarter Balanced vision and support the assessment system.

DELIVERING TECHNOLOGY-BASED ASSESSMENTS

Early in Year 3, Smarter Balanced hired a chief technology officer to lead its technology efforts. In Year 3, Smarter Balanced successfully administered the computer-administered pilot test to more 650,000 students across 5,000 schools. The pilot test provided an opportunity for the consortium to evaluate the functionality of accessibility features of the technology system, such as the use of pop-up glossaries. Based on the results of the pilot test, the delivery platform was enhanced prior to the field test. For example, the system added the capacity to integrate hand and machine scores for the field test and additional accessibility features. The field test in spring 2014 will provide a large-scale evaluation of the accessibility features and accommodations intended to be available when the assessment system is administered in the 2014-2015 school year. At the end of the grant period, Smarter Balanced will release the platform as an open-source platform.

The Smarter Balanced summative assessments will include a computer adaptive testing (CAT) algorithm, which means the system will adjust the difficulty of questions throughout the assessment based on the student's response to previous questions. This provides the summative assessments the ability to provide a more detailed measure of a student's knowledge and abilities with fewer questions than a traditional fixed-form assessment. In Year 3, the consortium made progress establishing the algorithm that will guide the CAT engine. In order to implement the adaptive component, Smarter Balanced needs data about each test item so that the test selects the appropriate next item for the student. Smarter Balanced has been conducting research, including running simulations, to evaluate the CAT design that is being developed to evaluate whether it is working as intended prior to the first operational administration in the 2014-2015 school year.

In Year 3, Smarter Balanced adopted a data privacy principle that guides its work; each member state retains control of its student-level data. Smarter Balanced began developing a data privacy policy consistent with this principle. In the past year, the consortium began developing a data warehouse and reporting system. These will become operational in Year 4 prior to the launch of the operational assessment system. Smarter Balanced developed a Beta system for reporting and solicited feedback from member states. Based on the feedback, Smarter Balanced has been working to revise and improve the system, including increasing the speed for loading reports. Smarter Balanced also continued developing its item authoring and item pool application. This will be used by the consortium to develop, review, and store items and tasks. This application will be used after the field test to store the items and tasks that have been developed as the consortium transition to the post-grant sustainable assessment system.

The consortium continues to rely on the expertise of the Architecture Review Board (ARB). The ARB is comprised of state elementary and secondary and higher education representatives and members of the Smarter Balanced technology work group. It meets twice per month to oversee further development of the IT systems architecture and makes recommendations regarding technology standards to the executive committee throughout the development of the assessment system.

DISTRICT AND SCHOOL TECHNOLOGY READINESS

During Year 3, Smarter Balanced released several tools to assist states in preparing their districts and schools for the implementation of the pilot, field test, and operational assessments. In fall 2012 (and

updated in November 2013), the consortium published the *Technology Strategy Framework and Testing Device Requirements*, which provides schools with the minimum hardware specifications and bandwidth calculations (see tables 1 and 2). The information provided assists schools and districts in evaluating which of their existing computers will support the administration of the consortium's assessment system, the network requirements, and other system requirements.

Table 1. Smarter Balanced technology specifications

Operating System	Minimum Smarter Balanced Requirements for Current Computers[1, 2, 3]	Recommended Smarter Balanced Minimum for New Purchases
Windows	Windows XP (service pack 3) Pentium 233 MHz processor 128 MB RAM 52 MB hard drive free space	Windows 7+ 1 GHz processor 1 GB RAM 80 GB hard drive or at least 1 GB of hard drive space available
Mac OS X	Mac OS X 10.4.4 Macintosh computer with Intel x86 or PowerPC G3 (300 MHz) processor, 256 MB RAM, 200 MB hard drive free space	Mac OS X 10.7+ 1 GHz processor 1 GB RAM 80 GB hard drive or at least 1 GB of hard drive space available
Linux	Linux (Ubuntu 9-10, Fedora 6) Pentium II or AMD K6-III 233 MHz processor 64 MB RAM, 52 MB hard drive free space	Linux (Ubuntu 11.10, Fedora 16) 1 GHz processor 1 GB RAM 80 GB hard drive or at least 1 GB of hard drive space available
iOS	iPads 2 running iOS6	iPads 3+ running iOS6
Android	Smarter Balanced-certified* Android-based tablets running Android 4.0+	Smarter Balanced-certified* Android-based tablets running Android 4.0+
Windows	Windows-based tablets running Windows 8+ (excluding Windows RT)	Windows-based tablets running Windows 8+ (excluding Windows RT)
Chrome OS	Chromebooks running Chrome OS (rolling release)	Chromebooks running Chrome OS (rolling release)

Source: www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Tech_Framework_Device_Requirements_11-1-13.pdf.

[1] The minimum Smarter Balanced requirements are generally equivalent to the minimum requirements of the associated eligible operating system. Users should refer to the minimum requirements of the operating system as a means of resolving any ambiguities in the minimum Smarter Balanced requirements.

[2] These guidelines do not supersede the minimum requirements of the operating systems.

[3] All hardware choices should consider the individual needs of students. Some students may need hardware that exceeds these minimum guidelines, and some students may require qualitatively different hardware. Tablets may require the use of a mouse.

* The Smarter Balanced "Device Certification Process" includes the certification for specific device models from manufacturers including, but not limited to, Android-based devices.

Smarter Balanced also released two additional tools to support states, districts, and schools in their planning for the consortium's assessment system. The first is an online tool for checking the bandwidth of a school's network. A school can identify the number of students who will be taking the test at one time to determine whether the school has sufficient network bandwidth. The second is a technology readiness

calculator. The calculator permits schools to identify the number of students and devices and the hours per day that the devices are available for testing at the school to determine the number of days and the associated network bandwidth required.

Table 2. Additional Requirements Applicable Across Operating Systems

Device Requirements	Minimum Smarter Balanced Requirements for Current Computers
Screen Size	10 inch class or larger with 1024 x 768 display resolution
Headphones/earphones	Available to students for use during the English language arts test and for students who require text-to-speech features on the mathematics test
Security	The device must have the administrative tools and capabilities to temporarily disable features, functionalities, and applications that could present a security risk during test administration.
Keyboards	External keyboards are required in all cases unless specified differently by a student's Individualized Education Program (IEP) or 504 plan. Any form of external keyboard that disables the on-screen virtual keyboard is acceptable. This includes mechanical, manual, plug and play, and wireless-based (e.g., Bluetooth, RF, IR) keyboards. The intent of this specification is to ensure the required display area is available to allow students to read multiple sources of complex item text and respond to source evidence for analytical purposes. While wireless keyboards are permissible, districts should be aware that high-density deployments of wireless keyboards and mice might interfere with each other or with the wireless network. Therefore, they should test the room configuration before the examination date and consider wired alternatives.
Pointing Device	A pointing device must be included. This may consist of a mouse, touch screen, touchpad, or other pointing device with which the student is familiar.
Form Factors	No restriction as long as the device meets the other stated requirements. These forms include desktops, laptops, netbooks, virtual desktops and thin clients*, tablets (iPad, Windows, Chromebooks, and Android), and hybrid laptop/tablets.
Network	Must connect to the Internet with a minimum of 20 Kbps available per student to be tested simultaneously. Local Web proxy caching servers are not recommended.

Source: www.smarterbalanced.org/wordpress/wp-content/uploads/2011/12/Tech_Framework_Device_Requirements_11-1-13.pdf.
NOTE: Minimum requirements represent a low compliance threshold. Ultimately, districts should attempt to exceed these requirements as many machines operating at these levels could struggle with sufficient on-board memory and processing to run secure browsers as well as other simultaneous running programs accumulated on the device over time.
*The resource (e.g., memory and processors) available to each client need to be equivalent or greater to the requirements for stand-alone hardware.

The consortium continued to make the technology readiness tool, developed in conjunction with PARCC during Year 2, available to its states. In Year 3, Smarter Balanced enabled the tool to report on gaps, if any, between the technology currently available and the technology necessary to administer the assessment system via computers in the 2014-2015 school year. Table 3 provides data from the technology readiness tool (and other sources used by the states, where applicable) for the number of local educational agencies (LEAs) that meet the consortium's minimum technology specifications to deliver the Smarter Balanced assessments.

Table 3. Technology readiness data, by state

State	Total number of LEAs in SY 2009-10	Number that submitted data on technology capacity	Number meeting consortium-defined specs
Alaska	133	0	DNP
California	984	562	446
Connecticut	187	194	142
Delaware	37	39	40
Hawaii	1	16	1
Idaho	138	135	DNP
Iowa	361	0	351
Maine	246	229	195
Michigan	791	770	285
Missouri	556	558	126
Montana	417	62	375
Nevada	17	9	15
New Hampshire	191	82	47
North Carolina	211	182	29
North Dakota	185	125	110
Oregon	197	17	197
Pennsylvania	634	0	DNP
South Carolina	86	80	43
South Dakota	156	97	140
Vermont	291	62	46
Washington	295	112	229
West Virginia	55	57	39
Wisconsin	442	142	340
Wyoming	48	19	28

Source: Data from the Smarter Balanced Annual Performance Report, submitted September 24, 2013.
Note: The Technology Readiness Tool (TRT) data may include school districts, administrative entities, correctional facilities, online schools, charter entities, and other out-of-district locations that may serve as testing centers. This definition may cause the reported number to be higher than the number of LEAs in an individual state.
*: Data from the Technology Readiness Tool.
DNP: Data not provided through the Technology Readiness Tool.
California: Approximately 15 percent of California schools completed the TRT. In an attempt to garner more information, the California Department of Education developed a brief survey to determine the preparedness of local educational agencies to administer the Smarter Balanced summative assessments. Respondents were asked to rate their level of confidence using a four point scale – completely confident, considerably confident, some level of confidence, and little confidence. Of the nearly 1,050 California school districts, 687 school districts responded to this survey. Of the 687 responding school districts, 446 (65 percent) indicated that, based on the Smarter Balanced minimum technology requirements, they are considerably to completely confident that if Smarter Balanced computer-assessments were administered today they would be able to complete testing within a 12-week window. The survey did not include responses from all districts, so this number is likely a little low. The responding local educational agencies serve approximately 87 percent of students enrolled in California public schools. Approximately 88 percent of responding school districts indicated they plan to use some of the $1.25 billion allocated in Assembly Bill 86 to purchase additional or new technology equipment in order to implement the Smarter Balanced assessments.
Connecticut: 142 districts represent the numbers of districts ready for online assessments that have over 70 percent of their devices meeting minimum requirements. There is no evidence that a school needs 100% of devices to meet minimum requirements, so a 70 percent threshold was set.
Delaware: Data from state testing experience.
Hawaii: Data from the Hawaii State Assessment; Hawaii is a single district state, within which are Complex Areas (16), managed by assistant superintendents.

Iowa: Data from the Iowa Condition of Education Report 2012; while the 2012 data indicates sufficient devices, the unresolved issue is sufficient bandwidth for online testing. Districts must reserve a portion of their bandwidth for critical systems (the network, student information systems, voice over IP telephones, email and other data systems used on a daily basis). Additionally, the pilot testing in Iowa revealed larger bandwidth consumption than we had anticipated.
Maine: Data from the Maine Learning Technology Initiative
Michigan: Data from the TRT. The number meeting consortium specifications was derived from the Device to Test-Taker Readiness Report; only districts at 100 percent were counted.
Missouri: LEAs are designated as having met the criteria for assessment technology readiness if they achieve a 100 percent rating on the TRT Device Indicator, Device to Test Taker Indicator, AND Network Indicator reports. Based on data self-reported in the TRT by LEAs, 83 percent have met the Device Indicator minimum requirements. In addition, 79 percent report to have met the Device to Test Taker Indicator goal and 88 percent report to have met the Network Indicator. While LEAs have made great strides in meeting the teaching and learning needs of the Generation Z student, 126 of 561 (22.46 percent) report meeting the criteria for all three indicators. Bandwidth and internal network readiness continues to be the primary challenge for LEAs.
Montana: Data from the TRT and State Education Agency Survey of local districts
Nevada: Data from Nevada Assessment Readiness Team Meetings
New Hampshire: Forty-seven represents the number of LEAs that have greater than 60 percent of their devices meeting the minimum requirements. Schools will be able to maximize the effective use of the devices that do meet the minimum requirements by establishing an appropriate number of testing days and associated number of testing sessions on each day that will allow for each student to use a compliant device.
North Carolina: Data from the TRT.
North Dakota: Data is a state-generated estimate, including data from the TRT and extrapolation.
Oregon: State has delivered math, English language arts, science, social science, and English language proficiency assessments online for over 10 years and is currently delivers these assessments to all students in the state. Oregon was the first to deliver a braille version of its adaptive test to blind and visually impaired students.
South Carolina: Established with Device to Test-Taker Indicators Report from the TRT.
South Dakota: Data from TRT, state technology inventory, and school communications.
Vermont: Data from the TRT; districts with over 50 percent devices meeting minimum requirements.
Washington: Data from combination of the TRT, state testing experience, and technology inventory.
West Virginia: Data from the TRT and state data collection.
Wisconsin: Data from School Speed Test Month and limited TRT data. Wisconsin based its readiness results primarily on over 63,000 results from Fall 2012 School Speed Test Month. During a 6-week period, 353 out of 447 districts, representing 1,281 out of 2,223 sites/schools performed at least 1 speed test. 33 percent of schools are Smarter-Balanced ready (>50 kbps/student); 43 percent are on the fence (10-20 to 50 kbps/student). Thus 76 percent meet or exceed Smarter Balanced bandwidth readiness already and 24 percent are below Smarter Balanced minimums of 10-20 kbps.
Wyoming: Data from the TRT and Wyoming Department of Education survey for state assessment program.

INTEROPERABILITY

The Smarter Balanced vision for the operational assessment, beginning in the 2014-2015 school year, calls for states to individually (or in small groups) contract with organizations to administer and score the assessments. In Year 3, Smarter Balanced released a request for information regarding the process to certify organizations to deliver the Smarter Balanced assessment system. The consortium also released initial item files for testing vendor input and test delivery components. Smarter Balanced met with technology providers, both organizations that develop devices and operating systems and organizations that develop assistive technologies, to address questions about how those systems will support the Smarter Balanced test delivery system. The consortium released technical specifications, test cases, and a website for organizations to evaluate their assistive technology. The consortium will evaluate and revise its system for determining interoperability following the field test in advance of the first operational administration in the 2014-2015 school year.

A related aspect of interoperability is specifying the technology standards for items and tasks so that the items can be rendered properly on assessment delivery system. Smarter Balanced continued during Year 3 to develop the technology standards for items and tasks. In November 2013, the consortium determined that it would develop two formats to support item development. The first, which the consortium has been using in item development, is the Smarter Balanced Assessment Item Format. It was developed by the consortium to match the Smarter Balanced item specifications and accessibility guidelines and has been used to deliver items for the pilot test in spring 2013 and will be used for the field test in spring 2014. The second is a set of compatible technical standards that are a profile of the Smarter Balanced Item Format

within the IMS Accessible Portable Item Protocol standards (APIP). This profile will enhance the current APIP standards to account for the newer, more complex item types. This will help support the Smarter Balanced vision by providing an option for states to contract with organizations that currently use the APIP format. In Year 4, the consortium will publish the technology standards.

Lessons Learned

States, districts, and schools continue to use the technology readiness tool to varying degrees to determine their readiness to administer the assessments. Smarter Balanced provided several additional tools to support states, districts, and schools. The technology readiness tool now provides a report indicating whether there are gaps in readiness. Additional tools test the network bandwidth and provide a calculator to determine whether there is sufficient number of devices and bandwidth in the school. In spite of these tools, however, the consortium and most member states do not have complete data on the number of schools or districts that are ready for all of their students to take the tests on computers in the 2014-2015 school year. The consortium will provide a paper-and-pencil version of the test for the first three years as a transition period.

Looking Ahead

The Smarter Balanced field test in spring 2014 will provide an opportunity for the consortium to try out its assessment delivery system with a larger number of students (more than 3.2 million) than the pilot test. The field test will include almost all of the functionality anticipated for the operational administration in the 2014-2015 school year, providing an opportunity to identify and resolve issues.

Following the field test, Smarter Balanced will make any necessary changes to the assessment delivery system and then release the technical specifications for the platform. In support of this and the Smarter Balanced Item Format discussed above, in Year 4, the consortium will launch a website dedicated to its technology standards (http://smarterapp.org) to create a community for the continued development and refinement of the open software that supports the Smarter Balanced assessment system. Smarter Balanced will also finalize its process to validate or demonstrate that organizations have the capability to administer the assessment system. Following the end of the field test, the consortium will also review the process to identify devices and assistive technologies that can be supported on the field test.

In addition, in Year 4 Smarter Balanced will continue to make available the technology readiness tool, the bandwidth tool, and the technology readiness calculator for states, districts, and schools to use to gather data and gauge their readiness to administer the assessments on computers in the 2014-2015 school year.

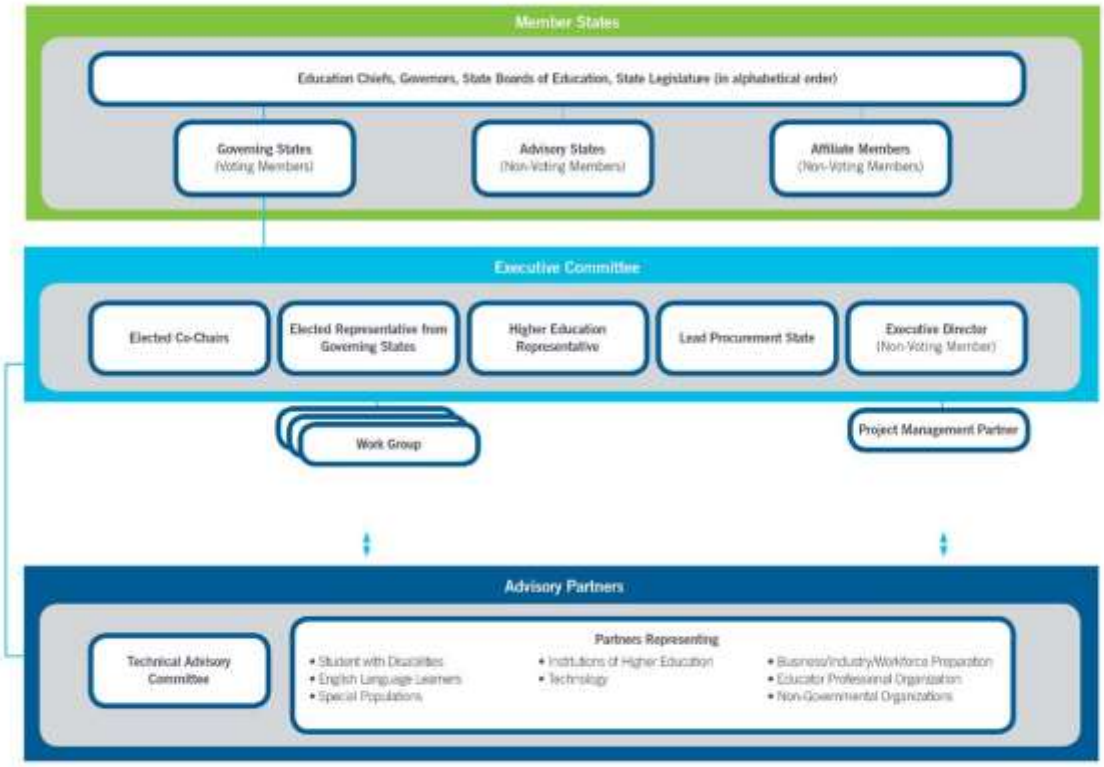
Finally, in Year 4 the consortium will complete all components of the technology system, including the item authoring and item pool application and the data warehouse and reporting application. The test delivery system that Smarter Balanced has been developing over the past three years and that was used for the pilot and field tests will be provided as an open-source resource for member states and any other interested party.

Governance

The consortium's approach to internal organization and capacity, project management, and procurement to permit timely decision-making and the efficient development and implementation of its assessment system.

The states that voluntarily joined Smarter Balanced continued to use the governance structure established in the first year of the grant to guide their decision making and establish policies to support the assessment system. Figure 4 shows the consortium's governance structure.

Figure 4. Smarter Balanced governance structure



Source: www.smarterbalanced.org/wordpress/wp-content/uploads/2012/02/Smarter-Balanced-Governance.pdf.

In Year 3, the consortium established policies related to the design of the assessments, initial ALDs and college content readiness, and technology to support the assessment system. Smarter Balanced states also voted in September to adopt a data privacy principle that affirms that each member state retains control of its student-level data and requires the consortium to develop a data privacy policy consistent with this principle. Contracts have been executed for most of the work necessary to develop the Smarter Balanced assessment system, including test development, test delivery, and reporting.

LEADERSHIP

In Year 3, Smarter Balanced maintained and refined its working governance structure under the continued general leadership of Executive Director Joe Willhoft and Chief Operating Officer Tony Alpert. The consortium hired a Chief Technology Officer, Director of Professional Learning, Director of State

Services, and Director of English Language Arts and Literacy. WestEd continues to serve as the Program Management Partner to support the consortium's work.

Smarter Balanced is led by an elected executive committee that meets biweekly and is comprised of nine voting members. The committee is led by two elected co-chairs. In Year 3, the co-chairs were Joseph Martineau of the Michigan Department of Education and Deb Sigman from the California Department of Education. The committee also consists of four at-large representatives from governing states, elected by K-12 leads in the member states; two higher education representatives, elected by higher education leads in member states; and one representative appointed by the state of Washington, the consortium's fiscal agent. In addition to the executive committee meetings, the consortium has twice-monthly meetings of all K-12 state leads, at which topics that require consortium-wide approval are discussed. The consortium continues to use workgroups and task forces made up of representatives from member states to help lead the consortium's work. To facilitate coordination across workgroups and task forces around design and implementation issues, and to ensure consistent understanding of the consortium's progress, Smarter Balanced continued the practice of convening twice-yearly collaboration conferences (in March and September 2013) that include workgroup co-chairs and contractors, as well as consortium leaders from Smarter Balanced and the states. These meetings included a public meeting with the chief state school officers in each governing state to discuss and vote on policy matters. In addition, the consortium continues to hold regular teleconferences and webinars with all chief state school officers and monthly teleconferences with the state leads for each member state to get their input and keep them apprised of the consortium's progress.

In Year 2, Smarter Balanced established a Sustainability Task Force. Members of the task force included the chief operating officer, executive director, one executive committee co-chair, three chief state school officers, two governors or their aides, two staff leading their state's involvement in the consortium, two state procurement experts, two state central information officers, and two higher education representatives. In Year 3, the task force completed its work and Smarter Balanced presented a sustainability plan which at the March 2013 collaboration conference. The member states voted to approve the post-grant sustainability plan and voted to co-locate with the Center for Research on Evaluation, Standards, and Student Testing (CRESST) at the Graduate School of Education and Information Studies at the University of California, Los Angeles (UCLA) at the end of the grant period in 2014. Smarter Balanced will remain a state-led organization with major budget, policy, and governance decisions resting with the governing states. Each member state will execute a memorandum of understanding (MOU) with UCLA to provide shared services to all member states (e.g., project management, maintenance of and enhancements to the assessment system, and on-going validity and other research studies). The member states will separately (either individually or in groups) contract with an organization to administer the Smarter Balanced assessments to their students beginning in the 2014-2015 school year.

PROCUREMENT

Smarter Balanced has awarded all of the significant contracts it expects to award during this grant. In Year 3, contracts were awarded for translating the assessments (see *Assessment Design and Development* for more information), pilot and field test scoring and field test item development, and field test administration. The request for proposals regarding the final contract Smarter Balanced anticipates awarding to support the development of its assessment system, for setting achievement standards, was released in December 2013.[9] As of December 31, 2013, the consortium had drawn down just more than half of its total grant award ($89,636,936.85).

[9] Smarter Balanced awarded a contract for these services in March 2014.

LESSONS LEARNED

During Year 2, Smarter Balanced contracted with individuals to serve as managers for the contracts related to the development of components of the assessment system. In Year 3, Smarter Balanced continued to have challenges receiving timely, high quality products from some contractors. Smarter Balanced leadership was actively engaged in raising concerns with contractors and requesting prompt resolution. In addition, as noted in *Assessment Design and Development*, Smarter Balanced substantially revised the item and task development process to address challenges identified in Year 2, instituting additional reviews and other quality control checks. As noted in *Professional Capacity, Communications, and Outreach*, Smarter Balanced identified challenges with the digital library resources and focused attention on making sure SLT and SNE members had sufficient training to identify, review, and approve resources that meet the consortium's expectations for quality.

In addition, in Year 3, Smarter Balanced realized the need to create ad hoc task forces and committees to address consortium needs, rather than relying primarily on the standing workgroups. The consortium created task forces or committees to define career readiness, sustainability, high school testing needs, and item packaging formats. In addition, Smarter Balanced established the FAAP, IQRP, and a finance committee to help guide its work.

LOOKING AHEAD

The consortium will continue to work closely among its member states (including both the K-12 and higher education leads for the state). Smarter Balanced will hold three collaboration conferences and two meetings of its TAC. Following the end of the field test in June 2014, Smarter Balanced will identify lessons learned from the field test and review and finalize policies related to test administration, accessibility, and test security. The consortium will also use the data from the field test to set initial achievement standards and to review the ALDs that were established in Year 3.

During Year 4, Smarter Balanced will complete its sustainability planning. States will negotiate and sign a MOU with UCLA that will define each of their respective roles and responsibilities..

Conclusion

Year 3 saw two significant milestones for Smarter Balanced. The consortium successfully conducted the pilot test to 650,000 students in more than 5,000 schools in spring 2013. Smarter Balanced also developed approximately 21,000 items and tasks for the field test that will take place in spring 2014. The successful pilot test and the release of practice tests in Year 3 provided an opportunity for states, districts, schools, teachers, and students to begin to see what the assessments will look like. The lessons learned from writing items for and conducting the pilot test and the results from the Department's Technical Review led the consortium to substantially revise its process for item development in Year 3 for the field test.

SUCCESSES

- *Assessment development*
  The pilot test provided an early opportunity to test the assessment delivery system and examine accessibility features. During Year 3, Smarter Balanced developed approximately 21,000 items and tasks in advance of the field test in spring 2014. The lessons learned from item development for the pilot test, many of which were also identified by the Department's Technical Review, led Smarter Balanced to significantly revise the item development process and institute additional quality control measures to improve the quality of the items and tasks developed for the field test.

  In Year 3, Smarter Balanced also developed accessibility guidelines and worked to establish the accessibility resources in advance of the field test so that the consortium, schools, and students can try them out during the field test. Smarter Balanced released online practice and training tests that include the consortium's accommodations and accessibility features so that states, stakeholders, and the public may try out the assessments. The consortium also adopted initial ALDs and drafted a career readiness framework.

- *Technology*
  Smarter Balanced has made significant progress realizing its vision for developing a technology-based assessment system. The consortium successfully administered the pilot test in Year 3 and prepared for a field test to more than 3 million students in Year 4. These activities provide opportunities for the consortium to pressure test the technology system in advance of the first full operational administration in the 2014-2015 school year. Smarter Balanced remains on track to release its assessment delivery platform as open-source software at the conclusion of the grant. This will support the consortium's sustainability plans as well as provide a benefit to all states, whether or not they are a member of Smarter Balanced. In Year 3, the consortium continued to support states, districts, and schools in their preparation for a technology-based assessment system by continuing to make the technology readiness tool available and including a gap analysis report that identifies whether the school or district has sufficient network bandwidth or devices. In addition, Smarter Balanced provided two tools to help school and district planning: an online tool for checking the bandwidth of a system's network and a technology readiness calculator to estimate the number of days and associated network bandwidth given the number of students, devices, and hours per day they are available for testing.

- *Sustainability*
  The consortium made significant progress in putting in place the structure for the sustainability of the consortium following the conclusion of the grant. The Smarter Balanced Sustainability Task Force developed a plan and the governing states, March 2013, voted to co-locate with CRESST at UCLA following the end of the grant period in 2014. Smarter Balanced has identified an approach to remain a state-led organization with major budget, policy, and governance decisions resting with the governing states.

CHALLENGES

- *Item development*
  As noted above, Smarter Balanced made significant progress during Year 3, developing approximately 21,000 items for the field test in spring 2014. Following the pilot test and the Department's Technical Review, Smarter Balanced increased its quality control measures by establishing the IQRP, revising item specifications, developing item quality criteria and an item audit process, and completely revising the mathematics performance tasks. The Department acknowledges the difficult work building a next-generation assessment system to measure whether students have the knowledge and skills necessary to succeed in college and the workforce. Smarter Balanced continued to experience challenges in Year 3 around adherence to established timelines and making sure the items and tasks developed met the consortium's quality criteria. The consortium should continue to evaluate whether its quality control processes are sufficient and provide close oversight over the development of future items and tasks to ensure that established timelines and quality criteria are being met. In Year 4, as Smarter Balanced develops additional items and tasks, it will need to be attentive to areas where it will need to improve the overall performance of the item pool.

- *Communications*
  The complexity of this project and the wide array of individuals that will be impacted by the consortium's assessment system means that it is vital for Smarter Balanced to continue, and expand, its communications with state leaders, educators, students, parents, members of the public, and key interest groups. Providing each audience a clear understanding about the assessment system is critical to the consortium's success. In Year 4, several consortium activities will provide an opportunity to demonstrate the benefits and importance of the assessment system. The field test, the launch of the digital library, and the initial achievement standards setting are each significant milestones for the consortium and an opportunity to engage the interested groups and the public in the consortium's work.

The Department is pleased to note that the consortium has identified and taken initial steps to mitigate these risks. Smarter Balanced is planning to increase its communications around the field test and achievement standards setting. In particular, Smarter Balanced will provide an opportunity for up to 250,000 educators, higher education faculty, parents, and other interested individuals from all member states to review items and data from the field test using an online tool to make recommendations for the level 3 cut score.

Throughout the remainder of Year 4, Smarter Balanced will continue to take steps toward developing its next-generation assessment system that will be implemented in the 2014-2015 school year. The consortium will:

- Administer and score the field test to more than 3.2 million students.
- Set initial achievement standards.
- Continue item and task development following the field test.
- Launch the digital library of formative tools and resources and continue identifying and developing resources to populate the library.
- Use data from the field test to identify the pool of items that will comprise the summative and interim assessments so that they are operational in the 2014-2015 school year.
- Continue to develop the technology system to support the development, administration, scoring, and reporting of the assessment system and establish the process to confirm adherence to the interoperability standards.
- Transition to UCLA to ensure the continued functioning of the consortium following the end of the grant.

Glossary

Accommodations means changes in the administration of an assessment, including but not limited to changes in assessment setting, scheduling, timing, presentation format, response mode, and combinations of these changes, that do not change the construct intended to be measured by the assessment or the meaning of the resulting scores. Accommodations must be used for equity in assessment and not provide advantage to students eligible to receive them.

Achievement level descriptors (ALDs) are text statements that articulate the knowledge, skills, and abilities represented at different levels of student performance. The levels of performance on the Smarter Balanced assessments are defined as below basic, basic, proficient, and advanced.

Achievement standard means the level of student achievement on summative assessments that indicates that (a) for the final high school summative assessments in mathematics or English language arts, a student is college- and career-ready; or (b) for summative assessments in mathematics or English language arts at a grade level other than the final high school summative assessments, a student is on track to being college- and career-ready. An achievement standard must be determined using empirical evidence over time.

The **American Recovery and Reinvestment Act of 2009 (ARRA)** was signed into law by President Obama on February 17, 2009. This historic legislation was designed to stimulate the economy, support job creation, and invest in critical sectors, including education. The U.S. Department of Education received a $97.4 billion appropriation.

College- and career-ready (or readiness) means, with respect to a student, that the student is prepared for success, without remediation, in credit-bearing, entry-level courses in an institution of higher education (IHE) (as defined in section 101(a) of the HEA), as demonstrated by an assessment score that meets or exceeds the achievement standard for the final high school summative assessment in mathematics or English language arts.

Common Core State Standards (CCSS) are K-12 English language arts and mathematics standards developed in collaboration with a variety of stakeholders including states, governors, chief state school officers, content experts, teachers, school administrators, and parents. The standards establish clear and consistent goals for learning that will prepare America's children for success in college and careers. As of January 2012, the Common Core State Standards were adopted by 45 states and the District of Columbia.

Common set of college- and career-ready standards means a set of academic content standards for grades K-12 that (a) define what a student must know and be able to do at each grade level; (b) if mastered, would ensure that the student is college- and career-ready by the time of high school graduation; and (c) are substantially identical across all states in a consortium. A state may supplement the common set of college-and career-ready standards with additional content standards, provided that the additional standards do not comprise more than 15 percent of the state's total standards for that content area.

Direct matriculation student means a student who entered college as a freshman within two years of graduating from high school.

English learner means a student who is an English learner as that term is defined by the consortium. The consortium must define the term in a manner that is uniform across member states and consistent with section 9101(25) of the ESEA.

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust on-going teaching and learning to improve students' achievement of intended instructional outcomes. Thus, it is done by the teacher in the classroom for the explicit purpose of diagnosing where students are in their learning, where gaps in knowledge and understanding exist, and how to help teachers and students improve student learning. The assessment is generally embedded within the learning activity and linked directly to the current unit of instruction. The assessments are typically small-scale (less than a class period) and short-cycle. Furthermore, the tasks presented may vary from one student to another depending on the teacher's judgement about the need for specific information about a student at a given point in time. Providing corrective feedback, modifying instruction to improve the student's understanding, or indicating areas of further instruction are essential aspects of a classroom formative assessment.

Governing state means a state that (a) is a member of only one consortium applying for a grant in the competition category, (b) has an active role in policy decision-making for the consortium, and (c) is committed to using the assessment system or program developed by the consortium.

Interim assessment is the term for the assessments that fall between formative and summative assessments. They typically evaluate students' knowledge and skills relative to a specific set of academic goals within a limited timeframe and are designed to inform decisions at both the classroom and school or district level. They may be given at the classroom level to provide information for the teacher, but unlike true formative assessments, the results of interim assessments can be meaningfully aggregated and reported at a broader level. As such, the timing of the administration is likely to be controlled by the school or district rather than by the teachers. They may serve a variety of purposes, including predicting a student's ability to succeed on a large-scale summative assessment, evaluating a particular educational program or pedagogy, or diagnosing gaps in a student's learning.

On track to being college- and career-ready means, with respect to a student, that the student is performing at or above grade level such that the student will be college- and career-ready by the time of high school graduation, as demonstrated by an assessment score that meets or exceeds the achievement standard for the student's grade level on a summative assessment in mathematics or English language arts.

The **Partnership for Assessment of Readiness for College and Careers (PARCC)** is one of two consortia of states awarded grants under the Race to the Top Assessment program to develop next-generation assessment systems that are aligned to common K-12 English language and mathematics standards and that will accurately measure student progress toward college and career readiness.

The **Smarter Balanced Assessment Consortium (Smarter Balanced)** is one of two consortia of states awarded grants under the Race to the Top Assessment program to develop next-generation assessment systems that are aligned to common K-12 English language and mathematic standards and that will accurately measure student progress toward college and career readiness.

A **student with a disability** means, for purposes of this competition, a student who has been identified as a student with a disability under the Individuals with Disabilities Education Act, as amended (IDEA), except for a student with a disability who is eligible to participate in alternate assessments based on alternate academic achievement standards consistent with 34 CFR 200.6(a)(2).

Summative assessments are generally given one time at the end of some unit of time such as the semester or school year to evaluate students' performance against a defined set of content standards. These assessments typically are given statewide and these days are usually used as part of an accountability program or to otherwise inform policy.

The Department of Education's mission is to promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access.

www.ed.gov